HfutEngine2009 Simulation 2D Team Description Paper

Li Shang, Long Li, Hao Wang, Baofu Fang

Hefei University of Technology, Hefei 230001, China

Abstract. This paper describes the background, the framework and the design feature of the HfutEngine2009. We put forward a new approach to do research on Multi Agent System. The method is based on mining teammate behavior. In this scene, an autonomous coach agent is able to get the current information of all teammates without noise, which can be modeled to compose patterns of teammates. At first coach agent gathers data from noisy environment to identify pattern of player agent. Then compute the probability of pattern compared with current situation by statistical calculations. Finally make best decision according to the result of mining teammate behavior.

1 Introduction

Team HfutEngine was founded in 2002 and took part in the RoboCup2002 of China. In the following years, HfutEngine is developing fast and joining many matches. In RoboCup China Open2007 we took the 2nd place of simulation soccer 2D. We took the 7th place of simulation soccer 2D in World RoboCup 2008. It is the second time for us to take part in The World RoboCup. We want to discuss Multi-Agent System and Robocup with anybody who is interested in them.

2 Framework of HfutEngine2D

In our exploitation we found that any Multi-Agent cooperation was based on how the single agent adapt to the Multi-Agent System. A system is steady when every element in the system can accommodate it. In that situation agent is not need to have single command. Our strategy is based on mining teammates' behaviors. Every Agent has its own evaluator to predict other teammate's behavior. Agent makes best decision by executor which based on the result of mining. The framework first founded in 2005. The former methods make use of evaluator to do action such as shoot, dribble and etc. In this way, the decision of the evaluator mainly depends on experience value, and player agent cannot cooperate smoothly. Now we do action based on the result of mining teammates' behaviors. The idea is not easy to perform. This structure is shown in Figure 1.

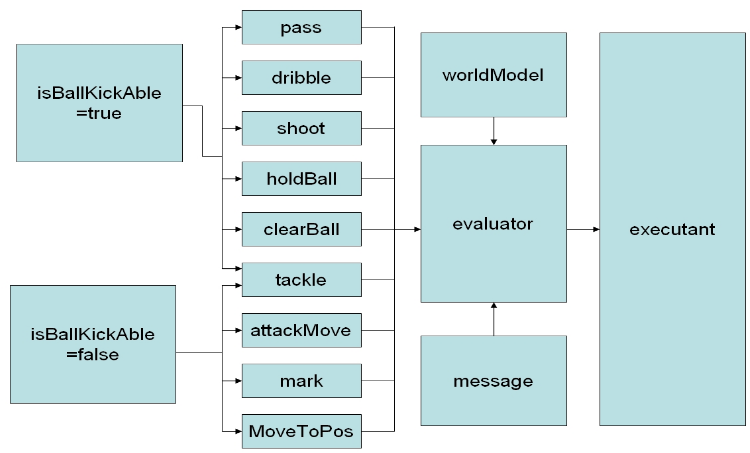

Fig. 1. The structure of HfutEngine2D.

3 High Level of HfutEngine2009

The HfutEngine2009's high decision includes two parts, evaluator and executor. Evaluator predicts the other teammate behavior by coach agent mining from environment, and we can make decision which action should be executed. Executor takes responsible for how to execute the action.

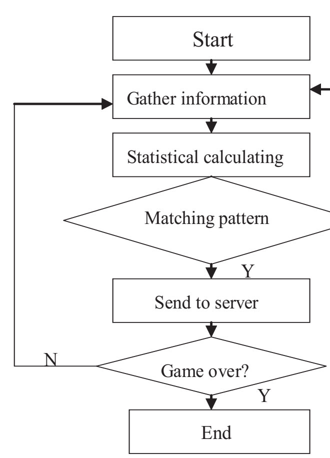

Fig. 2. Diagram of Mining Teammate Behavior

Firstly, online coach agent gathers information by visor from environment, then constructs model of teammates, using χ^2 test to identify other teammate patterns. Meanwhile, coach send new model to server during the game. Teammates get new model of others from server and use Q-learning to decide current measure. Finally, coach feeds new information back to surroundings. Figure 2 shows the diagram of mining teammate behavior.

3.1 Evaluator

The main application of evaluator is to make best decision. Because the environment is dynamic and stochastic, we can assume every action as random distribution. So we assume the decision of agent in the whole game as Gaussian distribution. Each agent makes decision independently. We all know that sum of Gaussian distribution is χ^{2} distributing. In this paper we make use of χ^{2} distributing to get the agent's pattern. For example, the ball is accelerated by agent A, and later received by agent B. This pattern can be considered as passing ball. Figure 3 shows pass ball pattern.

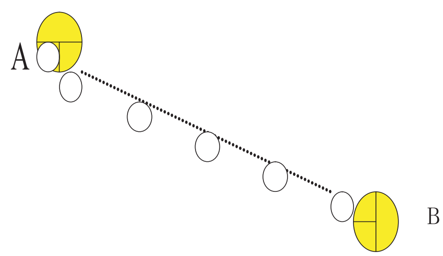

Fig. 3. Pass Ball Pattern

The step of evaluator can be described as follows:

Step1: We gather all the information of teammates from surroundings and check the current intention of agent.

Step2: We use following function to compute whether the current pattern is passing ball.

$$\chi^{2} = \sum_{i=1}^{k} \frac{(O_i - E_i)^2}{E_i}.$$

Attention: O_i means effect passing times, E_i means expect passing times

Step3: Use Q learning to train best Q table based on the pattern computing by step2.

Step4: Match passing ball pattern. We HfutEngine2D adopt formation523, and we take player2 pass ball for example, and the result shows in Table 1:

Step 5: Analyze the matching result. We assume $\alpha \equiv 0.1$, which means the ratio of pass ball is ≥ 0.9. We get 6.251 from distributing table. The computing

Table 1. passing ball result

Pass times	Num6	Num7	Num8	num10	Sum of pass times
effect passing times	93	454	172	181	1000
expect passing times	80	470	280	170	1000
Result	2.11	0.545	0.23	0.711	3.596

result is 3.596, we know that $3.596 \leq 6.251$, so we can confirm the feasibility of pass ball pattern is ≥ 0.9. Now we decide that the intension of agent is to pass ball. Figure 4 show the relationship between pass ball matching probability and training time.

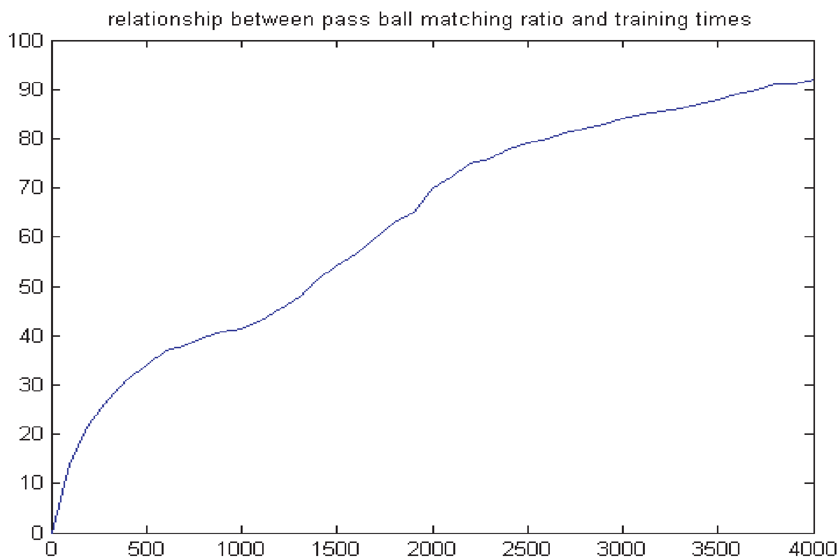

Fig. 4. Relationship Between Pass Ball Matching Probability and Training Time

3.2 Executor

The executor decides how the player performs. Some of them are gained by Machine Learning; others are deduced by the mathematic. Here take shoot for example. The action shoot is obtained by training. The training scene can be seen at Figure 5. The arithmetic is as follows:

Step 1:Fix the position of the ball and the shooting-player. Set the position of the goalie randomly and record the position.

Step 2:If the shooting-player takes the ball, then shoot to the fixed position in the goal.

Step 3:If the ball is kicked in the goal, then record the shoot is successful, whereas if the ball is hold up by the goalie then record the shoot is failure.

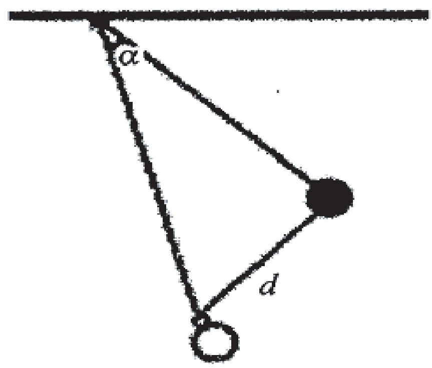

Fig. 5. Training for shoot

Repeat step 1 to step 3, we can get a set of value position of goalie, based on the value of the position of ball and position of Shooting, then we can compute the angle α and d. We use network of Radial Basis Function to train these values. Make the Gauss Nucleus Function and network framework as two inputs and one output. We have 3 parameters to learn, which are the center of Radial Basis Function, the weight from center-cell to output-cell, variance. The training result shows that the ratio of success increases to 91.28. Shoot-Advisor needs to submit method to the higher framework. The method needs 3 parameters.

4 Conclusion and Future Works

This paper proposed a new method to enhance cooperation in Multi Agent System. We use coach agent to mine useful information and teammate behavior to decide the current pattern of teammate.

The practice proves that the design idea of mining teammate behavior is very effective. The achievement that we have got in the past year further improves the ability of our team which makes a great leap. Table 2 shows the result of competing with some teams. It shows that HfutEngine2009 become much stronger.

We hope that we can improve more aspects of our team. We will further optimize this framework and exert its advantage and solve the problem of our

Table 2. the result of competing with some teams

Team	Ave Goals Scored	Ave Goals Conceded	win	draw	lose
HfutEngine2006	5.9	0.17	20	0	0
HfutEngine2007	4.4	0.15	20	0	0
HfutEngine2008	2.1	0.58	10	7	3
Mersad2006	2.25	0.7	15	3	2

team gradually. At the same time, we are planning to make a new way to forecast the situation in the playground based on experience which include the states of opponents and ball and so on. We also continue to make research on the Multi-Agent System and Machine Learning in order to enlarge the ratio of learning in our team. Meanwhile, the research will rather focus on the fast online learning than the online accumulated learning. The fast online learning make the player learn to change in time. The coming years, we will work hard to make a good result in the World RoboCup.

References

1. Zhulin AN, Jingjing YU, Hao WANG. Robocup Simulation League Goalie Design. st Proceedings of 1 Austria Open of RoboCup. 2003.
2. Xiaoping Chen, et al, Challenges in Research on Autonomous Robots, Communications of CCF, Vol. 3, No. 12, Dec 2007.
3. Kitano H, Tambe M,Stion P,et al. The Robocup synthetic agent challenge 97[A].Procceeding s of Fifteenth International Joint Conference on Artificial Intelligence (IJCA-97)[C]. Berlin:Springer,1997.24-29.
4. Baofu Fang,Hao Wang ,ect.The Survey of HfutEngine2005 Robot Simulation Soccer Team Design, Journal of Hefei University of Technology. Vol.29(9). Sep, 2006.
5. Gao Jian-qing, Wang Hao, Yu Lei. A Fuzzy Reinforcement Learning Algorithm and Application in RoboCup Environment. Computer Engineering and Applications,2006(6).
6. Yang LIU, Hao WANG, Baofu FANG, Hongliang Yao. Application of the method of supp vector regression in RoboCup. Journal of Hefei University of Technology Vol.30(10). Oct,2007.
7. Riedmiller, M., Gabel, T.: On Experiences in a Complex and Competitive Gaming Domain: Reinforcement Learning Meets RoboCup. In: Proceedings of the 3rd IEEE Symposium on Computational Intelligence and Games (CIG 2007).
8. Grefenstette, J., Schultz, A. (1994). An evolutionary approach to learning in robots. In Proceedings of the Machine Learning Workshop on Robot Learning, Eleventh International Conference on Machine Learning, New Brunswick, NJ.
9. P. Stone, R. S. Sutton, and G. Kuhlmann Reinforcement learning for RoboCup-soccer keep away Adaptive Behavior, 13(3):165–188, 2005.
10. Jinyi Yao,Jiang Chen,Zengqi Sun.An application in RoboCup combining Q-learning with Adversarial planning[D].Tsinghua University,2002.